Article

Current Biology

# Neural Correlates of the Conscious Perception of Visual Location Lie Outside Visual Cortex

## Highlights

- Perceived path of the double-drift illusion is not represented in visual cortex
- Consciously perceived location emerges in anterior areas beyond the visual cortex
- The illusion's long time constant matches properties of frontal not visual cortex
- Low-level features differentiate illusory stimuli in V2, V3, and MT+, but not V1

## Authors

Sirui Liu, Qing Yu, Peter U. Tse, Patrick Cavanagh

## Correspondence

sirui.liu.gr@dartmouth.edu

## In Brief

Using fMRI and MVPA, Liu and Yu et al. decode the perceived position of a moving stimulus with large perceptual displacement. They find that consciously perceived locations are not realized in the striate and extrastriate cortical areas traditionally associated with visual location processing but are instead represented in higher-order brain areas.

Liu et al., 2019, Current Biology *29*, 4036–4044
December 2, 2019 
https://doi.org/10.1016/j.cub.2019.10.033

CellPress

CellPress

Current Biology
Article

# Neural Correlates of the Conscious Perception of Visual Location Lie Outside Visual Cortex

Sirui Liu,[1,4,5,*] Qing Yu,[1,2,4] Peter U. Tse,[1] and Patrick Cavanagh[1,3]
[1]Department of Psychological and Brain Sciences, Dartmouth College, Hanover, NH 03755, USA
[2]Department of Psychiatry, University of Wisconsin–Madison, Madison, WI 53719, USA
[3]Department of Psychology, Glendon College, Toronto, ON M4N 3M6, Canada
[4]These authors contributed equally
[5]Lead Contact
*Correspondence: sirui.liu.gr@dartmouth.edu
https://doi.org/10.1016/j.cub.2019.10.033

## SUMMARY

**When perception differs from the physical stimulus, as it does for visual illusions and binocular rivalry, the opportunity arises to localize where perception emerges in the visual processing hierarchy. Representations prior to that stage differ from the eventual conscious percept even though they provide input to it. Here, we investigate where and how a remarkable misperception of position emerges in the brain. This "double-drift" illusion causes a dramatic mismatch between retinal and perceived location, producing a perceived motion path that can differ from its physical path by 45° or more. The deviations in the perceived trajectory can accumulate over at least a second, whereas other motion-induced position shifts accumulate over 80–100 ms before saturating. Using fMRI and multivariate pattern analysis, we find that the illusory path does not share activity patterns with a matched physical path in any early visual areas. In contrast, a whole-brain searchlight analysis reveals a shared representation in anterior regions of the brain. These higher-order areas would have the longer time constants required to accumulate the small moment-to-moment position offsets that presumably originate in early visual cortical areas and then transform these sensory inputs into a final conscious percept. The dissociation between perception and the activity in early sensory cortex suggests that consciously perceived position does not emerge in what is traditionally regarded as the visual system but instead emerges at a higher level.**

## INTRODUCTION

The representation of location is determined by an object's current retinal location in combination with several other sources of information, such as head and eye directions [1–4], eye movement plans [5], and the object's own motion [6]. Studies have shown that the visual system can predict the current location of a moving target by taking into account its velocity and the neural delays between the retina and the cortex [7]. It has been proposed that this predictive position shift, extrapolating the target ahead along its motion path, underlies several motion-induced position shift effects in which an object's location appears to be shifted by surrounding motion signals or by its own motion [6, 8–11].

The goal of the present study is to use predictive position shifts to investigate where the representation of perceived position emerges in the processing hierarchy. We used a probe that induces a remarkably large motion-induced position shift, namely, the double-drift illusion [12–15]. Compared with other well-known motion-induced position shift effects, this stimulus reveals an integration of motion signals over a second or more, leading to dramatic shifts in perceived position that can deviate from the physical motion trajectory by many degrees of visual angle (Figure 1A; Video S1). With such a long integration period, it is unlikely that early visual areas with their short integration time constants are responsible for the accumulation of position errors underlying this illusion. Thus, the double-drift stimulus presents the opportunity to explore where in the visual processing hierarchy position information transitions from retinally based, bottom-up encoding to high-level, motion-influenced perceptual representations associated with conscious visual experience. Specifically, if the patterns of neural activity that encode perception can be distinguished from those driven by the physical stimulus, we can identify the cortical areas where the percept first arises using multivariate pattern analysis (MVPA) of blood oxygen-level dependent (BOLD) signals measured using functional magnetic resonance imaging (fMRI) of the whole brain.

In two experiments, we demonstrate that the illusory motion paths of two double-drift stimuli with identical retinal but different perceived paths can be decoded from multiple brain regions, but the nature of the representation differs: the illusory paths are decodable in extrastriate areas but not in V1. However, cross-classification between the illusory paths and their matched physical paths in these areas showed no evidence that their patterns of response were related. In contrast, whole-brain cross-classification searchlight analyses reveal that higher-order brain areas share a common encoding of perceived position for the illusory and matched physical paths. Thus, our results indicate that different cortical regions are involved in representing different properties of the double-drift stimulus, with the early retinotopic visual areas possibly generating the local direction deviations driven by motion signals integrated over short durations, and

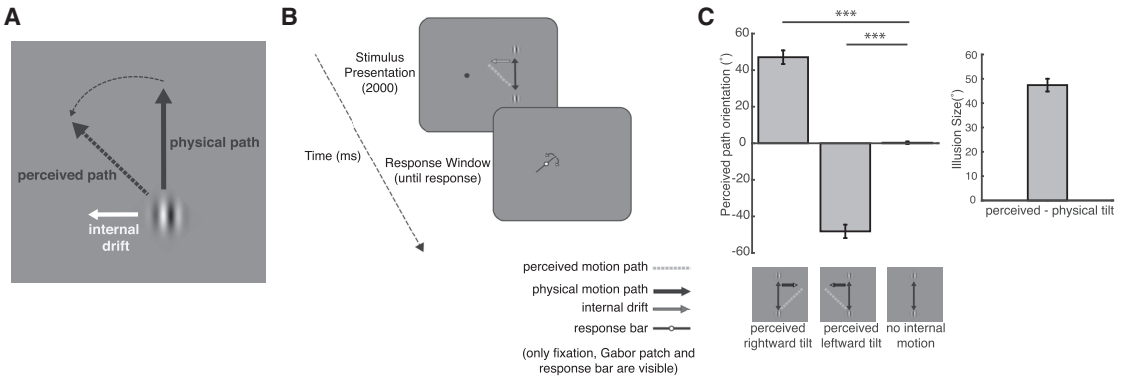


**Figure 1. Double-Drift Stimulus and Behavioral Task**

(A) Double-drift stimulus. A Gabor patch with vertical physical motion path can be perceived to be moving obliquely if its internal texture drifts orthogonally to its physical path. See also Video S1.

(B) Each trial began with a Gabor patch shown in the right hemifield moving vertically for 2 s, which then disappeared (example stimulus is a double-drift stimulus with a possible perceived motion path tilted leftward driven by its internal motion). A response bar then appeared at fixation and remained on the screen until participants adjusted its orientation to match the perceived motion path of the Gabor patch.

(C) Group-averaged perceived path orientation (°) of the double drift and control stimulus (no internal drift) and the illusion size of the double-drift effect. Error bars represent 95% CI, ***p < 0.001.

the higher-order regions possibly accumulating and storing position displacements based on extrapolations of those integrated motion directions to represent the long-lasting perceived motion path. An important implication of these results is that consciously perceived visual location is realized beyond the striate and extrastriate cortex traditionally associated with visual processing.

## RESULTS

### Experiment 1

#### *Perceived Path of the Double-Drift Stimulus Deviates from Its Physical Path*

We first conducted a behavioral task to measure the size of perceived position shift of the double-drift stimulus for each participant (see Figure 1B and STAR Methods for details). Consistent with previous literature, the perceived path orientation of the double-drift stimulus differed dramatically from that of a control stimulus that lacked internal motion (perceived rightward tilt: $p < 0.001$, Cohen's $d = 13.42$; perceived leftward tilt: $p < 0.001$, Cohen's $d = 14.89$; Figure 1C). Specifically, the perceived path orientation was biased toward its internal drift direction, suggesting that the motion-induced position shift of the double-drift stimuli was consistent for all subjects (average illusion size = 47.55° away from the veridical path orientation). There was no significant difference in the absolute amount of perceived direction shift between the two internal drift conditions (i.e., leftward versus rightward tilt) of the double-drift stimulus ($p = 0.60$, Cohen's $d = 0.30$).

#### *Perceived Paths Are Decodable in the Extrastriate Cortex But Do Not Share the Same Activation Patterns with Those of Matched Physical Paths*

We then used fMRI and MVPA to classify the activation patterns driven by two double-drift stimuli that moved along the same physical path but were perceived to be moving in opposite orientations driven by their different internal drift directions. Importantly, as the internal drift of the double-drift stimulus reverses its direction at the two endpoints of the motion path, both illusory trajectories have equal periods of leftward and rightward local motion across a complete back and forth cycle, so the only difference between the two conditions is their perceived motion direction. We compared these perceived motion paths with those of matched Gabor stimuli (lacking internal drift motion), which physically moved in the direction of the two illusory paths as measured in the behavioral task for each participant. During the scan, participants performed a contrast change detection task on the Gabor to ensure they were paying attention to it (see Figure 2A and STAR Methods for details). MVPA was first conducted in voxels that showed significantly greater BOLD responses to the motion path locations within early visual areas defined in a separate retinotopic mapping session. Figure 2B shows these 4 motion path regions of interest (MPROIs) for V1, V2, V3, and MT+ on a representative participant (see Table S1 for ROI sizes). When training and testing the linear support vector machine (SVM) classifier on the same stimulus conditions, classification accuracies for the two control stimuli (physical leftward versus rightward path orientation) were significantly above chance in all MPROIs ($p < 0.001$; p values were adjusted using the false discovery rate [FDR] method in this and all subsequent analyses; Figure 2C; see Table S1 for statistical results). This suggests that the activation patterns for the two stimuli with physical path orientations that matched those of the perceived path of the double-drift stimulus can be reliably differentiated in these early visual areas. However, classification accuracies for the two double-drift stimuli (illusory leftward versus rightward path orientation) were significantly above chance only in areas V2 ($p = 0.004$) and V3 ($p < 0.001$) but not in V1 ($p = 0.44$) or MT+ ($p = 0.12$) (Figure 2C; see Table S1 for statistical results). This indicates that activity patterns in V2 and V3 differed for the two perceptual motion paths of the double-drift stimulus, whereas activity patterns in V1 were indistinguishable for these two perceptual conditions. This is consistent with a recent study that finds evidence for the involvement of retinotopic areas with anatomically separated quadrantic representations of visual

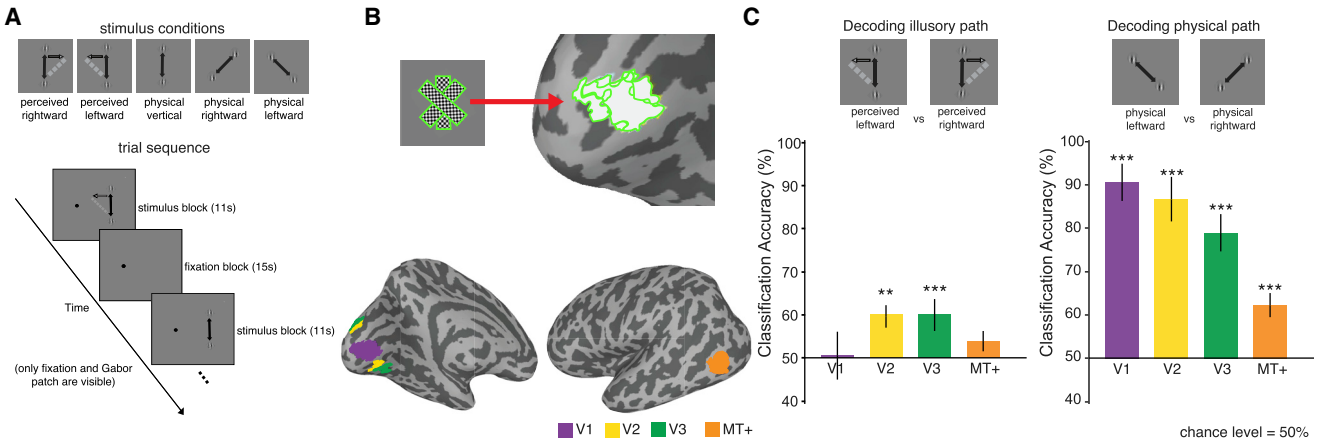


**Figure 2. Experiment 1: fMRI Trial Sequence and Decoding Accuracies in MPROIs**

(A) Five stimulus conditions in the main fMRI experiment: the double-drift stimuli had a vertical physical motion path with an orthogonal internal drift direction that could make the perceived motion path appear rotated either rightward or leftward relative to the physical path. The three control stimuli had either vertical, rightward, or leftward physical paths with no internal drift. Each trial was composed of 11 s of stimulus presentation followed by 15 s of fixation.

(B) Voxels for MPROIs were selected within each visual area by combining regions that showed greater activation for any of the three tilted rectangular checkerboard patterns (green outline not shown in experiment) than to fixation. Example ROIs are shown in a representative participant.

(C) Classification accuracies between the two double drift stimuli (perceived leftward versus rightward path) and the two control stimuli (physical leftward versus rightward path) in V1, V2, V3, and MT+.

Error bars represent ± 1 SEM, $^{**}p < 0.01$, $^{***}p < 0.001$.

See also Figure S1 and Table S1.

space, such as V2 and V3, in deriving the illusory motion of this stimulus [16].

To directly examine whether the activation patterns for the illusory paths of the double-drift stimulus share a similar structure with those of the matched physical paths, we conducted a cross-decoding analysis where we trained the linear SVM classifier on the data corresponding to the two double-drift stimuli and tested the classifier on the data corresponding to the two control stimuli. We also carried out the reverse analysis where we trained the classifier on the control stimuli and tested the classifier on the double-drift stimuli. Interestingly, classification accuracies from cross-classification in either direction were not significantly different from chance in any of the MPROIs ($p > 0.1$; see Table S1 for statistical results), including V2 and V3. Thus, although the activation patterns of the two illusory paths can be differentiated in V2 and V3, their representations carried different information from those of their matched physical paths in these two areas.

### *Representational Structure in Early Visual Areas Reveals Strongest Dissimilarity between Physically Different Paths*

Since the previous analysis suggested different representations of illusory and physical paths in early visual areas, we carried out a representational similarity analysis (RSA) to examine the representational structure of the five stimulus conditions in the MPROIs [17]. Figure 3 shows the dissimilarity matrices of the five stimulus conditions. Early visual areas V1–V3 exhibited the strongest dissimilarity between stimuli with different physical paths (physical leftward versus rightward versus vertical path) as compared to those that shared the same physical motion direction but with a large perceptual difference (double-drift stimuli: illusory leftward versus rightward; V1: $r = 0.91$, $p = 0.001$; V2: $r = 0.81$, $p = 0.016$; V3: $r = 0.70$, $p = 0.05$). This similarity structure confirmed that the representation of the double-drift stimulus in these areas was largely driven by its physical path. The representational structure in MT+, in contrast, showed high similarity between all stimulus conditions ($r = 0.27$, $p = 0.5$).

### *No Difference in BOLD Response Amplitude between Illusory and Matched Physical Paths*

We also calculated BOLD signal changes of each stimulus condition within each MPROI to examine whether the above-chance decoding accuracies for the illusory and matched physical paths could be detected at the univariate level. Group-averaged BOLD time courses of all stimulus conditions are shown in Figure S1. All MPROIs exhibited above-baseline activity for the five stimulus conditions ($p < 0.05$) except V1, which showed above-baseline activation only for the control stimulus with leftward motion direction. Importantly, we observed no difference in response magnitude for the two double-drift stimuli ($p > 0.1$) or for the two control stimuli in these MPROIs (*p*s > 0.1), suggesting that the above-chance decoding accuracies in these regions cannot be simply explained by differences at the aggregate activation level. In addition, there was no significant difference in mean signal intensity between the double-drift stimulus and the vertically moving control stimulus with no internal motion in any of these MPROIs (Figure S1C; $p > 0.1$), suggesting that the two conditions were matched in terms of stimulus energy. Thus, the failure to cross-decode in these regions was not simply due to a mismatch of internal motion in the double-drift and control stimulus.

### *Higher-Order Regions Show a Shared Representation of the Illusory and Matched Physical Paths*

To further explore brain areas that afford decoding of the illusory paths beyond our pre-defined visual ROIs, we conducted a whole-brain searchlight analysis using a 4-voxel-radius spherical searchlight. The same decoding analyses for the illusory and

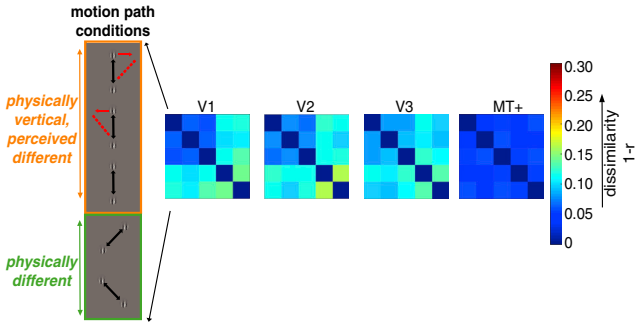


**Figure 3. Experiment 1: Representational Similarity Analysis**
Representational dissimilarity matrices for the five stimulus conditions in V1, V2, V3, and MT+.
See also Figure S4.

matched physical paths as for the ROI-based analysis were conducted, and results of the searchlight analysis were corrected using a cluster thresholding method for multiple testing (see STAR Methods for details). Figure 4 shows the group accuracy maps from the classification searchlight analysis for decoding the illusory paths (Figure 4A) and the matched physical paths (Figure 4B). We identified several clusters that showed above-chance decoding accuracies for the illusory path outside our pre-defined early visual ROIs: a large cluster that spanned the superior frontal and medial frontal gyrus and several clusters in the superior temporal gyrus, dorsal anterior cingulate (dACC) cortex, and the left postcentral gyrus. We also found a cluster spanning the early visual cortex that confirmed our significant decoding results in the ROI-based analysis (see Table S2 for a complete list of significant clusters). Decoding the matched physical paths yielded an even larger range of cortical regions, including visual, parietal, and frontal areas (Table S2).

Our ROI-based cross-decoding results showed that the activation pattern for the double-drift stimulus had little or no similarity to that of its matched physical path. It is, however, possible that a shared encoding of the illusory and matched physical path of the double-drift stimulus is represented somewhere outside our pre-defined visual ROIs. This shared encoding could be a marker of the emergence of the perceptual, as opposed to the retinal, location of the double-drift stimulus. We therefore conducted a whole-brain searchlight analysis using a cross-decoding classifier between the double-drift and control conditions to further explore the locus of any such shared representations. The results of this searchlight analysis should yield regions with similar patterns of activation for the double-drift and control stimulus that share the same perceived motion direction.

Interestingly, we found several clusters in anterior parts of the brain that showed above-chance cross-decoding between the illusory and matched physical paths (Figure 4C; see Table S2 for a complete list of significant clusters) but, in agreement with the previous ROI analysis, none in early visual areas. Specifically, we found clusters that showed above-chance cross-decoding in both directions (i.e., trained on double-drift then tested on control stimuli and vice versa) in the anterior cingulate and medial frontal gyrus in both hemispheres, the anterior part of the middle frontal gyrus in the left hemisphere, and the left inferior parietal lobule and parahippocampal gyrus. Besides these overlapping regions, cross-decoding from double-drift to control stimuli resulted in additional significant clusters in the middle and inferior frontal gyrus and medial frontal gyrus in both hemispheres; cross-decoding from control to double-drift stimuli produced additional significant clusters in the right precentral gyrus and left parahippocampal gyrus. In addition to these cortical clusters, we also found several subcortical clusters as detailed in Table S2.

To exclude the possibility that the failure in cross-decoding in regions such as the early visual cortex was caused by (potentially) subtle difference in mean signal intensity across conditions, even though their differences in mean activation amplitude were not statistically significant (Figure S1), we performed an additional searchlight analysis where we removed the grand mean of each stimulus condition within each searchlight. The results remained qualitatively similar to those of the original cross-decoding searchlight analysis as shown in Figure 4C, with no significant clusters observed in early visual cortex (see Figure S2 and Table S3). This suggests that failure to cross-decode in regions such as the early visual cortex was not simply due to a difference in stimulus-driven responses between the stimulus conditions of the training (e.g., physical leftward versus rightward paths) and testing datasets (e.g., illusory leftward versus rightward paths).

## Experiment 2

### *Perceived Paths Do Not Share the Same Activation Patterns in Early Visual Areas with Those of Matched Physical Paths Using Counterphase-Flickering Gabor Patches*

In experiment 1, we showed that the double-drift and the control stimulus were matched in terms of the amplitudes of stimulus-driven responses. However, given that early visual cortex is sensitive to stimulus-specific, low-level feature changes, it is still possible that the failure in cross-decoding in these regions was due to a mismatch of internal motion in the double-drift (4 Hz) and the control stimulus (0 Hz). Thus, we conducted a second experiment using a control stimulus to match the motion energy induced by the internal motion of the double-drift stimulus. Similar to the one used in experiment 1, this control stimulus was a Gabor patch moving in the perceived path of the double-drift stimulus but with its internal grating counterphase flickering at 4 Hz. Importantly, the perceived path direction of the counterphase-flickering control was not significantly different from its physical motion direction ($p = 0.17$), suggesting that adding the counterphase flicker did not affect the perceived path of the physical motion direction (Figure 5A). Moreover, there was no significant difference in mean BOLD signal intensity between the double-drift stimulus and this new control stimulus in any of the MPROIs (Figure S3C; $p > 0.1$), suggesting that including the counterphase flicker did not result in changes at the univariate level and that the new control stimulus matched the double-drift stimulus in terms of stimulus energy. Beside this new control stimulus, all experimental and analysis procedures were identical to those in experiment 1 (see STAR Methods for details). Thus, if it was indeed the difference in internal motion of the illusory and control stimuli that degraded the cross-classification in early visual areas, cross-classification in these regions would be successful with this control

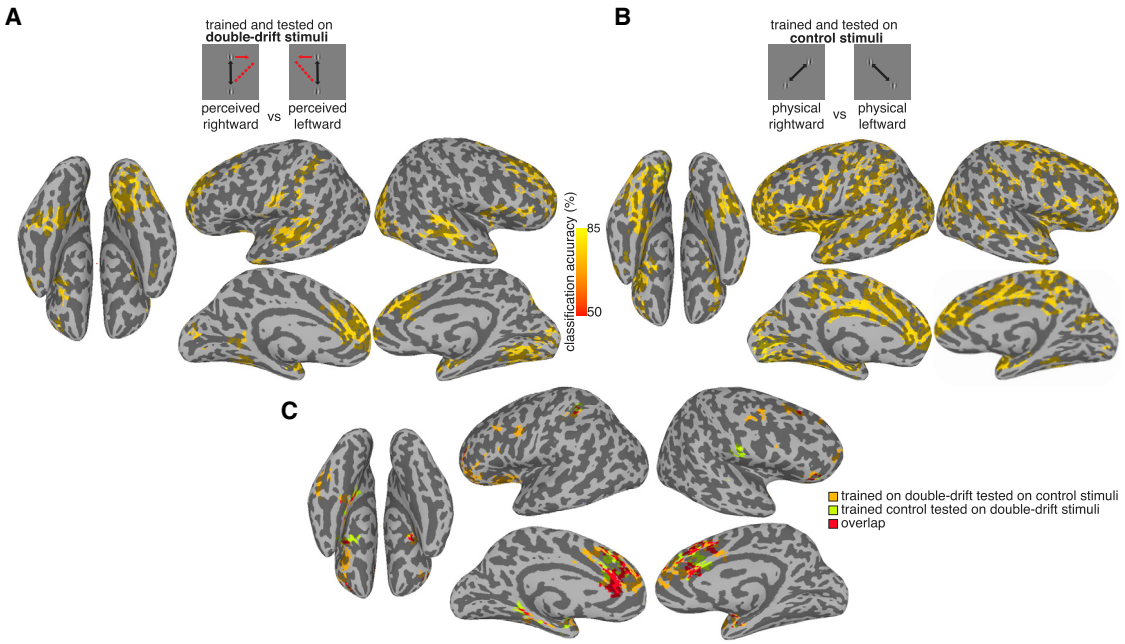


**Figure 4. Experiment 1: Searchlight Maps with Significant Above-Chance within- and Cross-Condition Classification Accuracy**

(A) Areas supporting significant decoding of the illusory paths.

(B) Areas supporting significant decoding of the matched physical paths.

(C) Areas supporting significant cross-decoding of illusory and matched physical paths. Orange represents significant clusters when training the classifier on double-drift and testing on the control stimuli. Green represents significant clusters when training the classifier on the control and testing on the double-drift stimuli.

Results were thresholded at $p = 0.01$ and FDR-corrected across clusters at $p < 0.05$.

See also Figure S2, Table S2, and Table S3.

stimulus, since it has a better match for the motion energy of the double-drift stimulus.

Overall, results in experiment 2 largely replicated those of experiment 1. Specifically, in the ROI-based MVPA analysis, classification between the two illusory paths (average illusion size = 39.2°; Figure 5A) was only successful in extrastriate visual areas V3 ($p < 0.001$) and MT+ ($p = 0.006$), but not in V1 ($p = 0.16$) or V2 ($p = 0.095$). Decoding between the physical paths of the counterphase-flickering control was successful in all MPROIs ($p < 0.001$; Figure 5B; see Table S1 for statistical results). Since the same double-drift stimuli were employed in both experiments 1 and 2, we also combined the results for decoding the illusory paths from the two experiments ($n = 19$) to increase statistical power. Results showed that classification accuracies were significantly above chance in V2 ($p = 0.001$), V3 ($p < 0.001$), and MT+ ($p = 0.001$) but not in V1 ($p = 0.22$; Figure 5C; see Table S1 for statistical results). Moreover, consistent with the results from experiment 1, cross-decoding between the illusory paths of the double-drift stimulus and the matched physical paths of the counterphase-flickering control stimuli failed in all MPROIs ($p > 0.1$; see Table S1 for statistical results). These null results confirmed that failure in cross-decoding in experiment 1 in these ROIs was not due to low-level feature mismatches between the double-drift and static Gabor control stimuli, since the new control stimulus that yielded similar results had a 4 Hz flicker that matched the internal motion energy of the illusory stimulus.

Consistent with the results in experiment 1, representational dissimilarity matrices showed that V1–V3 exhibited the strongest dissimilarity between stimuli with different physical paths of the counter-flickering control compared to those that shared the same physical motion direction but differed perceptually (Figure S4; V1: $r = 0.80$, $p = 0.01$; V2: $r = 0.79$, $p = 0.013$; V3: $r = 0.66$, $p = 0.05$). The representational structure in MT+ again showed high similarity between all stimulus conditions ($r = -0.51$, $p = 0.13$).

In the whole-brain searchlight analysis, decoding for the illusory (Figure 6A) and physical paths (Figure 6B) both yielded extended clusters across multiple cortical regions that were comparable to those found in experiment 1 (see Table S4 for a complete list of significant clusters). For the cross-decoding searchlight analysis, despite being conducted in different scanners and with different control stimuli, we found that the two critical aspects from experiment 1 still held up in experiment 2: the successful cross-decoding between the illusory and matched physical paths was observed only in anterior parts of the brain (especially in the lateral and medial frontal cortex), while no successful cross-decoding was found in the early visual cortex. Specifically, we found significant above-chance clusters in both directions (i.e., trained on double-drift then tested on control stimuli and vice versa) in the dorsolateral prefrontal cortex in both hemispheres and the left medial frontal gyrus. Besides these overlapping regions, cross-decoding from double-drift to the new control stimuli resulted in additional significant clusters in

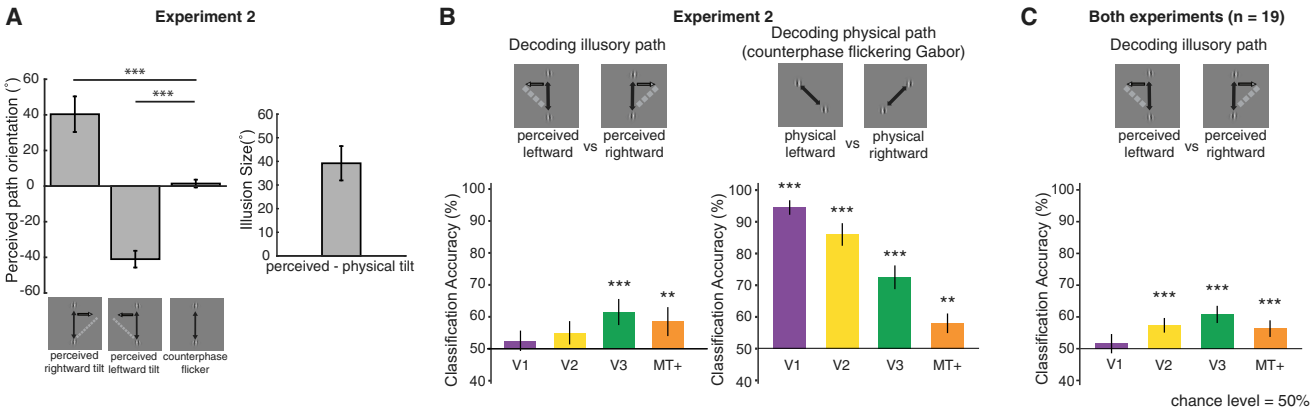


**Figure 5. Experiment 2: Behavioral Performance and Decoding Accuracies in MPROIs**

(A) Group-averaged perceived path orientation (°) of the double-drift stimulus and control stimulus (counterphase-flickering Gabor) and the illusion size of the double-drift effect. Error bars represent 95% CI.

(B) Classification accuracies for the two double-drift stimuli (perceived leftward versus rightward path) and the two counterphase-flickering control stimuli (physical leftward versus rightward path) in V1, V2, V3, and MT+.

(C) Both experiments (n = 19): classification accuracies for the two double-drift stimuli (perceived leftward versus rightward path).

Error bars represent ± 1 SEM, $^{**}p < 0.01$, $^{***}p < 0.001$.

See also Figure S3 and Table S1.

the middle frontal gyrus in both hemispheres and the left angular gyrus; cross-decoding from the new control to double-drift stimuli produced additional significant clusters in the left medial frontal gyrus, temporo-parietal junction and superior temporal gyrus (Figure 6C; see Table S4 for a complete list of significant clusters). These searchlight results confirmed that the failure in cross-decoding in experiment 1 in early visual cortex was not due to a mismatch of internal motion between the double-drift and control stimuli but a difference between the representations of the illusory and matched physical paths in these regions.

## DISCUSSION

Here, we localized cortical areas associated with perceived object positions when they differed from positions registered on the retina. Normally, the location of retinal input and the corresponding perceived locations are well matched, but, in a remarkable motion-induced position shift effect, the double-drift illusion, there is a dramatic mismatch between physical and perceived positions, and the large perceptual displacement can build for a second or more [12, 13].

Although we found that the perceived path of the double-drift stimulus can be decoded in extrastriate visual areas, results from the cross-classification analyses revealed that the activation patterns that differentiated the illusory paths in these early visual regions were not related to those that encoded the matched physical paths. Therefore, the basis of our classification results in these early regions is likely unrelated to the perceived path per se but might arise from lower-level properties of the stimulus such as the combined vector of the local and global motion [13]. For example, leftward internal drift was associated with an upward external motion in one case (e.g., double-drift stimulus with perceived leftward trajectory) and with a downward external motion in the other (e.g., double-drift stimulus with perceived rightward trajectory).

Interestingly, the significant cross-classification clusters found in our searchlight analyses were primarily in anterior parts of the brain, such as the lateral prefrontal cortex (LPFC), dACC (the cingulo-opercular control network), and medial prefrontal cortex (MPFC), that are known to be involved in executive control [18–22] and working-memory-related processing [23–26]. This indicates that the neural coding for the perceived path in these higher-order regions was driven by a representation of the illusory motion path that was similar enough to that of the matched physical path to permit cross-decoding. In comparison to representations in higher cortical areas, those in early visual cortex are sensitive to local, fine-grained feature-level changes and therefore may not permit such cross-classification. Indeed, our results show that this high-level representation of perceived as opposed to real stimulus positions is not shared with or projected back down to early visual areas. Similarly, experiments on binocular rivalry also identified structures in LPFC, among others, as representing the dominant perceived image almost exclusively, whereas early visual cortex represented both eyes' image, and their strengths were modulated to some extent by the perceptual state [27, 28]. Thus, the observed above-chance classification in extrastriate visual areas for the perceived path of double-drift stimulus might suggest that these areas encode the combined local motion signals integrated over short durations. These local direction errors are the base data that get integrated into the illusory path but likely do not account for the illusion alone, because these errors appear to accumulate over long durations, and cells in these early processing stages do not have the second-long integration time windows that the double-drift stimulus requires to build up its position deviations. In contrast, other motion-induced position shifts effects such as the flash-grab stimulus only integrate motion signals over about 90 ms [8, 29]. Given the short decay time constants for orientation cells in early visual areas [30], it is possible that a motion-position integration process of such a long duration requires higher-order brain areas

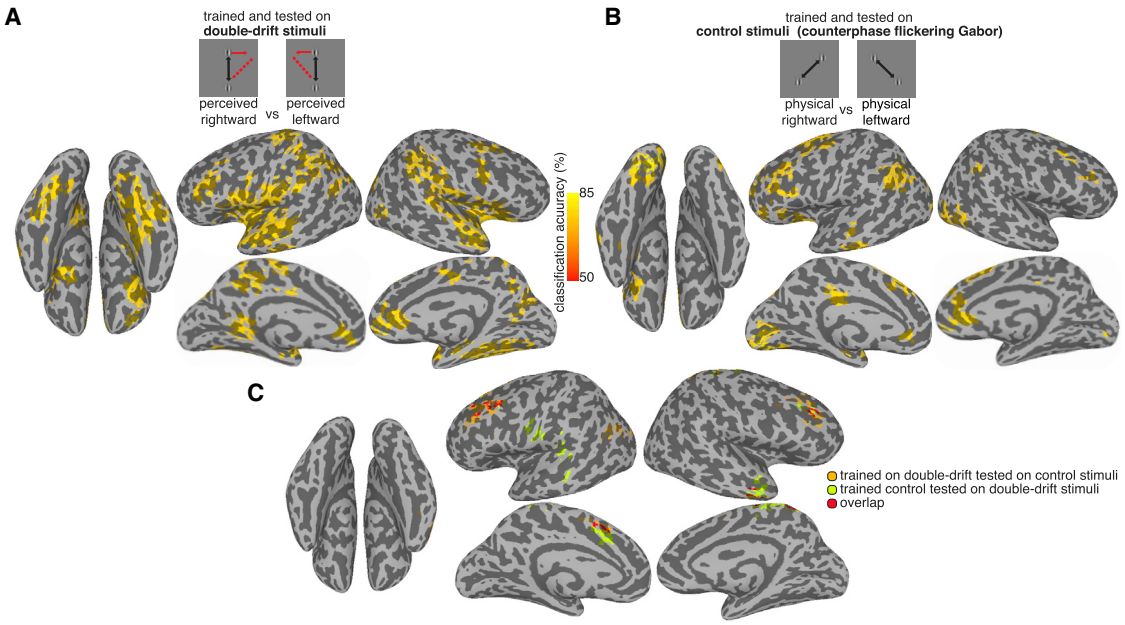


**Figure 6. Experiment 2: Searchlight Maps with Significant Above-Chance within- and Cross-Condition Classification Accuracy**

(A) Areas supporting significant decoding of the illusory paths.

(B) Areas supporting significant decoding of the matched physical paths.

(C) Areas supporting significant cross-decoding of illusory and matched physical paths. Orange represents significant clusters when training the classifier on double-drift and tested on the control stimuli. Green represents significant clusters when training the classifier on the control and tested on the double-drift stimuli.

Results were thresholded at $p = 0.01$ and FDR corrected across clusters at $p < 0.05$.

See also Table S4.

to store and accumulate position offsets in order to form a consistent motion trajectory. Thus, our results suggest that the higher-order areas where we find significant cross-classification could be candidate areas that accumulate outputs from the extrastriate cortex so that perception continues to drift farther away from the real path for over a second. Since these higher-order regions have been implicated in transforming sensory representation to different representational formats to serve different functional purposes [31–34], successful cross-classification in these areas may reflect similar representational changes from a sensory format to a different, more abstract format, which could allow for generalization between different physical stimuli in a shared format of perceptual experience. Note that the observed successful cross-decoding in higher-order areas cannot be simply explained by similarity in decision making or motor responses for the illusory and matched physical paths, because participants performed an orthogonal task (reporting a 50% reduction of the Gabor's contrast) in the scanner that did not involve decisions or motor responses that corresponded to the direction of motions.

Our finding that there is no shared activity in early visual cortex between the physical and perceived paths conflicts with previous fMRI studies of motion-induced position shifts [35–37]. In the case of the "flash-grab" illusion [8], it was shown that neural activity for the perceived position shifts correlates with that of perceptually matched physical stimuli in V1 through V3 but not in higher-order areas [37]. Similarly, for the "flash-drag" illusion [38], activity in MT+ also shows strong correlation between perceived and matched physical positions [36]. There are many differences between these studies and ours, but we speculate that the involvement of top-down attentional signals may account for the discrepancies. It is well established that saccades, which are closely linked to spatial attention [39, 40], are directed to the perceived rather than physical locations for the motion-induced position shifts that displace the target along the direction of motion, such as the flash-lag and flash-drag stimuli [41–44]. For the double-drift illusion, however, saccades are directed to the physical location of the stimulus, rather than its perceived location [12]. Given that the dissociation of saccades and perception is unique to the double-drift stimulus, it is possible that attentional shifts, like saccades, are not affected by this illusion either, so that any downward projections from areas involved in attentional shifting circuitry would prioritize its physical locations, rather than its perceived, illusory ones. If this assumption is correct, this stimulus affords the unique possibility of probing where the perceptual coordinates of object position arise in the processing hierarchy without the confound of top-down attentional projections.

In summary, our data reveal cortical areas that process visual stimuli in perceived as opposed to retinotopic coordinates. We found that the representation of perceived position likely emerges much later in the processing hierarchy than the early visual areas, even if extrastriate areas provide the instantaneous direction errors that are then integrated in later areas to compute the final percept. There are many aspects of visual consciousness such as perceived color, brightness, duration, shape,

direction, as well as the binding of these features. Although our data do not address the cortical areas where all these perceived features emerge, they do place clear constraints on the neural correlates of the perceived position at which all of these aspects are experienced, at least for this particular type of stimulus. Surprisingly, the construction of perceived location appears to emerge in areas beyond the visual cortex.

## STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- LEAD CONTACT AND MATERIALS AVAILABILITY
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Subjects
- METHOD DETAILS
  - Stimuli
  - Pre-scan behavioral task
  - MRI acquisitions
  - Main experiment runs
  - Region-of-interest localization runs
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Behavioral data analysis
  - fMRI data analysis
- DATA AND CODE AVAILABILITY

### SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.cub.2019.10.033.

### ACKNOWLEDGMENTS

This research was funded by NSF grant 1632738 (P.U.T. and P.C.) and an NSERC Canada grant (P.C.).

### AUTHOR CONTRIBUTIONS

All authors designed the experiment and wrote the manuscript. S.L. and Q.Y. conducted the experiment and analyzed the data. Q.Y. contributed unpublished analytic tools. P.U.T. and P.C. supervised the entire project.

### DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: June 2, 2019
Revised: September 23, 2019
Accepted: October 17, 2019
Published: November 21, 2019

### REFERENCES

1. Bremmer, F., Ilg, U.J., Thiele, A., Distler, C., and Hoffmann, K.P. (1997). Eye position effects in monkey cortex. I. Visual and pursuit-related activity in extrastriate areas MT and MST. J. Neurophysiol. *77*, 944–961.
2. Brotchie, P.R., Andersen, R.A., Snyder, L.H., and Goodman, S.J. (1995). Head position signals used by parietal neurons to encode locations of visual stimuli. Nature *375*, 232–235.
3. DeSouza, J.F.X., Dukelow, S.P., and Vilis, T. (2002). Eye position signals modulate early dorsal and ventral visual areas. Cereb. Cortex *12*, 991–997.
4. Snyder, L.H., Grieve, K.L., Brotchie, P., and Andersen, R.A. (1998). Separate body- and world-referenced representations of visual space in parietal cortex. Nature *394*, 887–891.
5. Duhamel, J.R., Colby, C.L., and Goldberg, M.E. (1992). The updating of the representation of visual space in parietal cortex by intended eye movements. Science *255*, 90–92.
6. Whitney, D. (2002). The influence of visual motion on perceived position. Trends Cogn. Sci. *6*, 211–216.
7. Nijhawan, R. (1994). Motion extrapolation in catching. Nature *370*, 256–257.
8. Cavanagh, P., and Anstis, S. (2013). The flash grab effect. Vision Res. *91*, 8–20.
9. De Valois, R.L., and De Valois, K.K. (1991). Vernier acuity with stationary moving Gabors. Vision Res. *31*, 1619–1626.
10. Nijhawan, R. (2002). Neural delays, visual motion and the flash-lag effect. Trends Cogn. Sci. *6*, 387–393.
11. Nijhawan, R. (2008). Visual prediction: psychophysics and neurophysiology of compensation for time delays. Behav. Brain Sci. *31*, 179–198, discussion 198–239.
12. Lisi, M., and Cavanagh, P. (2015). Dissociation between the perceptual and saccadic localization of moving objects. Curr. Biol. *25*, 2535–2540.
13. Tse, P.U., and Hsieh, P.J. (2006). The infinite regress illusion reveals faulty integration of local and global motion signals. Vision Res. *46*, 3881–3885.
14. Kwon, O.S., Tadin, D., and Knill, D.C. (2015). Unifying account of visual motion and position perception. Proc. Natl. Acad. Sci. USA *112*, 8142–8147.
15. Shapiro, A., Lu, Z.L., Huang, C.B., Knight, E., and Ennis, R. (2010). Transitions between central and peripheral vision create spatial/temporal distortions: a hypothesis concerning the perceived break of the curveball. PLoS ONE *5*, e13296.
16. Liu, S., Tse, P.U., and Cavanagh, P. (2018). Meridian interference reveals neural locus of motion-induced position shifts. J. Neurophysiol. *119*, 2091–2099.
17. Kriegeskorte, N., Mur, M., and Bandettini, P. (2008). Representational similarity analysis - connecting the branches of systems neuroscience. Front. Syst. Neurosci. *2*, 4.
18. Wagner, A.D., Maril, A., Bjork, R.A., and Schacter, D.L. (2001). Prefrontal contributions to executive control: fMRI evidence for functional distinctions within lateral Prefrontal cortex. Neuroimage *14*, 1337–1347.
19. Koechlin, E., Ody, C., and Kouneiher, F. (2003). The architecture of cognitive control in the human prefrontal cortex. Science *302*, 1181–1185.
20. Dosenbach, N.U.F., Fair, D.A., Miezin, F.M., Cohen, A.L., Wenger, K.K., Dosenbach, R.A.T., Fox, M.D., Snyder, A.Z., Vincent, J.L., Raichle, M.E., et al. (2007). Distinct brain networks for adaptive and stable task control in humans. Proc. Natl. Acad. Sci. USA *104*, 11073–11078.
21. Dosenbach, N.U.F., Fair, D.A., Cohen, A.L., Schlaggar, B.L., and Petersen, S.E. (2008). A dual-networks architecture of top-down control. Trends Cogn. Sci. *12*, 99–105.
22. Ridderinkhof, K.R., Ullsperger, M., Crone, E.A., and Nieuwenhuis, S. (2004). The role of the medial frontal cortex in cognitive control. Science *306*, 443–447.
23. Curtis, C.E., and D'Esposito, M. (2003). Persistent activity in the prefrontal cortex during working memory. Trends Cogn. Sci. *7*, 415–423.
24. Ester, E.F., Sprague, T.C., and Serences, J.T. (2015). Parietal and frontal cortex encode stimulus-specific mnemonic representations during visual working memory. Neuron *87*, 893–905.
25. Ungerleider, L.G., Courtney, S.M., and Haxby, J.V. (1998). A neural system for human visual working memory. Proc. Natl. Acad. Sci. USA *95*, 883–890.
26. Yu, Q., and Shim, W.M. (2017). Occipital, parietal, and frontal cortices selectively maintain task-relevant features of multi-feature objects in visual working memory. Neuroimage *157*, 97–107.

27. Leopold, D.A., and Logothetis, N.K. (1996). Activity changes in early visual cortex reflect monkeys' percepts during binocular rivalry. Nature *379*, 549–553.
28. Panagiotaropoulos, T.I., Deco, G., Kapoor, V., and Logothetis, N.K. (2012). Neuronal discharges and gamma oscillations explicitly reflect visual consciousness in the lateral prefrontal cortex. Neuron *74*, 924–935.
29. Chung, S.T.L., Patel, S.S., Bedell, H.E., and Yilmaz, O. (2007). Spatial and temporal properties of the illusory motion-induced position shift for drifting stimuli. Vision Res. *47*, 231–243.
30. McLelland, D., Baker, P.M., Ahmed, B., and Bair, W. (2010). Neuronal responses during and after the presentation of static visual stimuli in macaque primary visual cortex. J. Neurosci. *30*, 12619–12631.
31. Badre, D., Kayser, A.S., and D'Esposito, M. (2010). Frontal cortex and the discovery of abstract action rules. Neuron *66*, 315–326.
32. Badre, D., and Nee, D.E. (2018). Frontal cortex and the hierarchical control of behavior. Trends Cogn. Sci. *22*, 170–188.
33. Wallis, G., Stokes, M., Cousijn, H., Woolrich, M., and Nobre, A.C. (2015). Frontoparietal and cingulo-opercular networks play dissociable roles in control of working memory. J. Cogn. Neurosci. *27*, 2019–2034.
34. Ploran, E.J., Nelson, S.M., Velanova, K., Donaldson, D.I., Petersen, S.E., and Wheeler, M.E. (2007). Evidence accumulation and the moment of recognition: dissociating perceptual recognition processes using fMRI. J. Neurosci. *27*, 11912–11924.
35. Whitney, D., Goltz, H.C., Thomas, C.G., Gati, J.S., Menon, R.S., and Goodale, M.A. (2003). Flexible retinotopy: motion-dependent position coding in the visual cortex. Science *302*, 878–881.
36. Maus, G.W., Fischer, J., and Whitney, D. (2013). Motion-dependent representation of space in area MT+. Neuron *78*, 554–562.
37. Kohler, P.J., Cavanagh, P., and Tse, P.U. (2017). Motion-induced position shifts activate early visual cortex. Front. Neurosci. *11*, 168.
38. Whitney, D., and Cavanagh, P. (2000). Motion distorts visual space: shifting the perceived position of remote stationary objects. Nat. Neurosci. *3*, 954–959.
39. Moore, T., and Zirnsak, M. (2017). Neural mechanisms of selective visual attention. Annu. Rev. Psychol. *68*, 47–72.
40. Awh, E., Armstrong, K.M., and Moore, T. (2006). Visual and oculomotor selection: links, causes and implications for spatial attention. Trends Cogn. Sci. *10*, 124–130.
41. de'Sperati, C., and Baud-Bovy, G. (2008). Blind saccades: an asynchrony between seeing and looking. J. Neurosci. *28*, 4317–4321.
42. Schafer, R.J., and Moore, T. (2007). Attention governs action in the primate frontal eye field. Neuron *56*, 541–551.
43. Zimmermann, E., Morrone, M.C., and Burr, D. (2012). Visual motion distorts visual and motor space. J. Vis. *12*, 10.
44. van Heusden, E., Rolfs, M., Cavanagh, P., and Hogendoorn, H. (2018). Motion extrapolation for eye movements predicts perceived motion-induced position shifts. J. Neurosci. *38*, 8243–8250.
45. Mathworks. (2012). MATLAB and Statistics Toolbox Release 2012b (Mathworks).
46. Cox, R.W. (1996). AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. Comput. Biomed. Res. *29*, 162–173.
47. Fischl, B., Sereno, M.I., and Dale, A.M. (1999). Cortical surface-based analysis. II: Inflation, flattening, and a surface-based coordinate system. Neuroimage *9*, 195–207.
48. Kleiner, M., Brainard, D.H., Pelli, D., Ingling, A., Murray, R., and Broussard, C. (2007). What's new in Psychtoolbox-3. Perception *36*, 1–16.
49. R Core Team (2013). R: A Language and Environment for Statistical Computing (R Foundation for Statistical Computing). https://www.R-project.org/.
50. Allaire, J. (2012). RStudio: Integrated Development Environment for R (RStudio).
51. Hanke, M., Halchenko, Y.O., Sederberg, P.B., Hanson, S.J., Haxby, J.V., and Pollmann, S. (2009). PyMVPA: a python toolbox for multivariate pattern analysis of fMRI data. Neuroinformatics *7*, 37–53.
52. Engel, S.A., Rumelhart, D.E., Wandell, B.A., Lee, A.T., Glover, G.H., Chichilnisky, E.J., and Shadlen, M.N. (1994). fMRI of human visual cortex. Nature *369*, 525.
53. Sereno, M.I., Dale, A.M., Reppas, J.B., Kwong, K.K., Belliveau, J.W., Brady, T.J., Rosen, B.R., and Tootell, R.B. (1995). Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. Science *268*, 889–893.
54. Huk, A.C., Dougherty, R.F., and Heeger, D.J. (2002). Retinotopy and functional subdivision of human areas MT and MST. J. Neurosci. *22*, 7195–7205.
55. Talairach, J., and Tournoux, P. (1988). Co-planar Stereotaxic Atlas of the Human Brain (Thieme Verlag).
56. Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. R. Stat. Soc. B *57*, 289–300.
57. Stelzer, J., Chen, Y., and Turner, R. (2013). Statistical inference and multiple testing correction in classification-based multi-voxel pattern analysis (MVPA): random permutations and cluster size control. Neuroimage *65*, 69–82.

# STAR★METHODS

## KEY RESOURCES TABLE

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---|---|---|
| Software and Algorithms | | |
| MATLAB 2015a | MathWorks [45] | https://www.mathworks.com/ |
| AFNI | [46] | https://afni.nimh.nih.gov/ |
| FreeSurfer | [47] | http://surfer.nmr.mgh.harvard.edu |
| Psychophysics Toolbox | [48] | https://psychtoolbox.com/ |
| R version 3.3.2, RStudio interface | [49, 50] | https://www.r-project.com/ |
| PyMVPA | [51] | http://www.pymvpa.org/ |

## LEAD CONTACT AND MATERIALS AVAILABILITY

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Sirui Liu (Sirui.Liu.gr@dartmouth.edu). This study did not generate new unique reagents.

## EXPERIMENTAL MODEL AND SUBJECT DETAILS

### Subjects

Nine individuals from the Dartmouth College community participated in Experiment 1 of this study (5 females; age range: 21-32, mean age = 26.6 +- 3.1) and ten new individuals from the Dartmouth College community participated in Experiment 2 of this study (6 females; age range: 18-32, mean age = 21 +- 4.5). All participants were naive to the purpose of this study and had normal or corrected-to-normal vision. Written, informed consent approved by the Committee for the Protection of Human Subjects at Dartmouth College was obtained from each participant prior to each experimental session. Participants were screened by the Dartmouth Brain Imagining Center fMRI Subject Safety Screening Sheet and received a compensation of $20/hour.

## METHOD DETAILS

### Stimuli

All stimuli were generated using MATLAB 2015a [45] and PsychToolbox-3 [48]. The stimulus in the behavioral and the main fMRI experiments was based on the double-drift stimulus used in previous literature [12] (see Figure 1A and Video S1). We used a Gabor pattern (sinusoidal grating within a Gaussian envelope) with a spatial frequency of 0.5 cycle/dva (cycles per degree of visual angle) and 100% contrast presented on a uniform gray background (53 cd/m$^2$). The standard deviation of the contrast envelope was 0.4 degree of visual angle [dva]. The Gabor pattern moved back and forth along a linear path of length 5 dva, with a speed of 5 dva/sec (external motion). The sinusoidal grating had the same orientation as the motion path, and drifted in the orthogonal direction with a temporal frequency of 4 Hz (internal motion) while reversing its direction at the two endpoints every 1 s ('double-drift stimulus'), or stayed static (Experiment 1: 'control stimulus') or counterphase-flickered at 4 Hz (Experiment 2: 'control stimulus') during the trial. The midpoint of the trajectory was placed at 5 dva to the right of the screen center. A black fixation point (0.3 dva diameter) was presented at 3 dva horizontal to the left of the screen center throughout all the behavioral and MRI experiments. We moved the fixation to this location so that our stimulus was 8 dva away from fixation. This was the eccentricity at which previous research [12] found a large perceptual effect. In the pre-scan behavioral task, participants reported the perceived orientation of the motion path using a black line ('response bar') centered at fixation that was 0.05 dva in width and 5 dva in length.

### Pre-scan behavioral task

Stimuli were generated using an Apple iMac Intel Core i5 (Cupertino, CA) and were displayed in a dark room on a 16'' ViewSonic G73f CRT monitor (1024 × 768 pixels at 90-Hz) placed 57-cm from the participant with their head stabilized on a chinrest during the experiment. Figure 1B shows a sample trial of the pre-scan behavioral task. Participants were instructed to keep their gaze at the fixation point throughout the experiment. In each trial, a Gabor patch was shown in the periphery and moved back and forth along a vertical path for 2 s. Its internal texture drifted either leftward or rightward, or remained static (Experiment 1) or counterphase-flickered (Experiment 2). For each participant, the drift condition of the internal texture was randomized across trials. Following the stimulus, participants were instructed to rotate the response bar by pressing the corresponding keyboard keys (up arrow for counterclockwise; down arrow for clockwise) until its direction matched the perceived angle of the motion trajectory of the Gabor. The response bar was

presented at a random orientation for each trial and remained on the screen until participants were satisfied with their response and pressed the space bar for the next trial. Overall, each participant completed ten adjustment trials for each internal drift condition for a total of 30 trials that lasted about 15 minutes. The measured path orientation was then used to create the control and localizer stimuli in the fMRI experiment for each participant (see Behavioral data analysis for details).

### MRI acquisitions

The scanning in Experiment 1 was conducted in a 3T MRI scanner (Philips Intera Achieva) with a 32-channel head coil at the Dartmouth Brain Imaging Center at Dartmouth College. For each subject, we collected functional BOLD activity using an echo-planar imaging (EPI) sequence (TR = 2 s, TE = 35 ms, voxel size = 3 × 3 × 3 mm, flip angle = 90°, FOV = 240 × 240 mm) and a high resolution anatomical scan using T1-weighted MPRAGE sequence at the end of each scanning session (voxel size = 1 × 1 × 1 mm).

The scanning in Experiment 2 was conducted in a 3T MRI scanner (Siemens PRISMA) with a 32-channel head coil at the Dartmouth Brain Imaging Center at Dartmouth College. For each subject, we collected functional BOLD activity using an echo-planar imaging (EPI) sequence (TR = 2 s, TE = 32 ms, voxel size = 3 × 3 × 3 mm, flip angle = 90°, FOV = 240 × 240 mm) and a high resolution anatomical scan using T1-weighted MPRAGE sequence at the end of each scanning session (voxel size = 1 × 1 × 1 mm).

### Main experiment runs

Stimuli were presented on a screen (Experiment 1: 47.5 cm width; Experiment 2: 35.6 cm width) at the back of the scanner through an LCD projector. The screen resolution was 1024 × 768 pixels with 60 Hz refresh rate. The projected stimuli were viewed through a mirror located on the head coil with a viewing distance of 101.6 cm (Experiment 1) and 124.5 cm (Experiment 2). Participants completed 10 fMRI main experimental runs. In each run, after an initial 4 s blank fixation period, participants viewed a total of fifteen trials, each of which was composed of a 11 s stimulus block followed by a 15 s fixation period. Each stimulus block was composed of a Gabor patch presented in the right hemifield that moved back and forth along a linear path for five repetitions (2 s each repetition) and then disappeared for 250 ms in between repetitions. Five stimulus conditions were presented in each run, including two double-drift stimuli that had the same vertical physical motion path but with opposite internal drift directions that could make the perceived motion path appear rotated either rightward or leftward relative to the physical motion path, and three control stimuli that had either vertical, rightward or leftward physical motion paths with no internal drift (Experiment 1) or with 4 Hz counterphase flicker (Experiment 2). The path orientations of the control stimuli were individually calculated from the responses in the pre-scan behavioral task for each participant. Figure 2A illustrates the five stimulus conditions and a sample trial sequence for the main fMRI experiment. To make sure participants were attending to the stimulus, the contrast of the presented Gabor stimulus reduced 50% randomly in each run for 200 ms and participants were asked to press a response button each time they saw the change. Participants viewed three blocks per stimulus condition in each run, and the order of the trials was randomized for each run. In total, each experimental run was 394 s long.

We also conducted two additional EPI runs using a rectangular checkerboard pattern flickering at 8 Hz that covered the spatial extent of the perceived and physical motion paths of the double-drift stimulus. The checkerboard pattern was centered at 8 dva horizontal to the right of the fixation with its height the length of the motion path of the Gabor pattern and its width the size of the Gabor stimulus. Figure 2B shows the three conditions (vertical, leftward or rightward tilted rectangular checkerboard pattern) for the stimulus location localizer runs. The two oblique checkerboard stimuli were tilted in the direction of the perceived motion path for the double-drift stimulus. The tilt angle was individually calculated from the responses in the pre-scan behavioral task for each subject. Each run contained an initial 4 s fixation block and fifteen trials, each of which was composed of a 10 s stimulus block with a flickering checkerboard pattern followed by a 12 s. blank fixation period. There were five trials per stimulus condition for a total of fifteen trials per run with the order of the blocks randomized for each participant. To maintain fixation, participants were asked to press a response button each time they saw the color of the fixation point changed.

### Region-of-interest localization runs

In addition to the main experiment, participants completed a separate scanning session that included a standard retinotopic mapping procedure and three MT+ localizer runs (292 s each). We followed the standard retinotopic mapping procedure [52, 53] by using clockwise or counterclockwise rotating checkerboard wedges (flickering at 4 Hz, ten 192 s runs) to map polar angle and using expanding or contracting checkerboard rings (flickering at 4 Hz, four 192 s runs) to map eccentricity. The fixation point in our experiment was moved 3 dva horizontal to the left of the center of the screen to match that of our main experiment. MT+ was functionally localized for each participant following the procedure from previous literature [54]. In each run, participants viewed seven 16 s blank fixation periods interleaved with six 30 s stimulus blocks. The stimulus was composed of one hundred 0.3-dva diameter black dots spanning the whole visual field that moved coherently, or flickered at 30 Hz, or remained static on the screen. Each of the three stimulus conditions was presented twice in each run with the order randomized for each subject. For the coherently moving condition, the dots could be moving rightward, leftward, vertically upward or downward, expanding, contracting, rotating clockwise or counterclockwise at 7 dva/s with 100% coherence, while resetting their locations every 367 ms. To make sure they were fixating, participants were asked to press a response button each time they saw the color of the fixation point changed during all localizer runs.

## QUANTIFICATION AND STATISTICAL ANALYSIS

### Behavioral data analysis

For each participant, we first determined the perceived path of the double-drift stimulus (with leftward or rightward perceived motion paths) and of the control stimulus that lacked internal motion (Experiment 1) or with counterphase flicker (Experiment 2), by measuring the perceived angle away from the physical, vertical orientation. One sample t tests were conducted using R and RStudio v.1.0.136 to compare the mean differences between the perceived path orientation of the control stimulus with vertical path orientation [49, 50]. Paired-samples t tests were conducted to compare the mean differences of the perceived path orientation between the double-drift stimulus and the control stimulus. Significance of the tests were determined at $p < 0.05$. The magnitude of the double-drift illusion was then calculated individually by taking the difference between these two measurements. A positive value of the illusion size indicates that the perceived motion orientation was biased toward that of the internal drift. The average of this value was then used in the following scanning session as the motion direction for the control stimuli that moved obliquely with no internal drift as well as the tilt angle for the rectangular checkerboard pattern in the localizer runs to define ROIs for the motion path of the stimulus for each subject.

### fMRI data analysis

#### *Preprocessing*

Functional imaging data was preprocessed using AFNI [46]. For each participant, the EPIs were first registered to the last run of each scan session and then motion corrected, linearly detrended, and z-scored within each run. The anatomical images were aligned to the EPI scans of the same session. Localizer data were further smoothed with a 4-mm FWHM Gaussian kernel. For the searchlight analysis, the EPI scans were normalized to the Talairach standard space [55].

#### *ROI definition*

The cortical surface of each participant was first reconstructed with FreeSurfer [47] using the high-resolution anatomical images from the localizer session. All data in the localizer runs were first mapped onto this cortical surface to define the ROIs. Early visual areas left V1, V2, and V3 were individually drawn by hand on individual surfaces based on the phase angle and eccentricity maps computed from data in the retinotopic mapping session. MT+ was individually defined on surface based on data from the MT+ localizer runs using beta coefficient values calculated from a General Linear Model (GLM) analysis that specified voxels that responded more strongly to moving than to stationary dot patterns ($p < 10^{-4}$ after correcting for multiple tests using false-discovery rate (FDR) [56]. To identify the voxels that responded to the motion path of the double-drift and control stimulus within each of the ROIs, we then selected voxels that showed significantly greater activation for any of the three tilted rectangular checkerboard patterns than to fixation ($p < 10^{-4}$, FDR corrected) in the left hemisphere and only these voxels were included for the rest of the ROI-based analyses. All these surface-defined ROIs were then individually mapped back into the volume space and aligned to the EPI data of the first scanning session by aligning the anatomical scans of the two sessions for subsequent analysis.

#### *Time course of BOLD activity*

To create the time series of BOLD activity change in each ROI, we averaged BOLD activity in all voxels within the ROI and calculated the percent signal change relative to baseline for each TR of each trial. Baseline was defined as the activity of the first TR of each trial. Average BOLD signal change at each time point was then calculated by averaging the percent signal change across trials within each condition. One-sample t tests against 0 were used to assess statistical significance above baseline for each TR and for each stimulus condition within each ROI at $p < 0.05$ after correcting for multiple tests using FDR [56]. In addition, paired samples t tests were used to compare BOLD activity between 1) the two double-drift stimulus conditions, 2) the two control stimulus conditions, and 3) the double-drift stimulus conditions and the vertically moving control stimulus at each time point within each ROI, and significance of the tests were determined at $p < 0.05$ after correcting for multiple tests using FDR [56].

#### *Multivariate pattern analysis (MVPA)*

All the subsequent analyses were performed using the PyMVPA toolbox [51]. We first performed MVPA within each ROIs. For each trial, we extracted raw data averaged for 6 to 14 s after trial onset (considering a 6 s of hemodynamic delay) and fed the averaged data into linear support vector machines (SVMs) to implement classification of stimulus conditions. We performed two types of classification analyses: the first analysis was to classify between the two physically different motion paths of the control stimulus, or between the two perceptually-different motion paths of the double-drift stimulus, using a leave-one-run-out cross-validation procedure. To examine whether the activation patterns of the double-drift stimuli resembled that of the corresponding control stimuli, a second cross-decoding analysis was conducted using the same data, except that the training and test data were from separate conditions (i.e., training with the data corresponding to the control stimuli with matched physical motion path and testing with the data corresponding to the double-drift stimuli with physically vertical but perceived different motion path; and vice versa). Statistical significance of classification accuracies across subjects for each individual experiment and for the two experiments combined was determined by randomly shuffling the stimulus condition labels 1000 times to construct null distributions for each ROI and testing for significance above chance at $p < 0.05$ after correcting for multiple tests using false-discovery rate (FDR) [56].

The SVMs were further combined with a spherical searchlight procedure for whole-brain classification analysis. Specifically, we applied a volume-based searchlight analysis by sliding a 4-voxel-radius spherical linear SVM classifier voxel-by-voxel over the whole brain. As with the ROI-based analysis, the searchlight analysis was performed for decoding illusory paths, the matched physical paths, and cross-decoding using a leave-one-run-out cross-validation procedure. Group-level statistical significance for the

searchlight analyses for each individual experiment was determined following a cluster thresholding approach [57]: 100 permuted searchlight accuracy maps were first generated for each subject by randomly permuting the stimulus condition labels across trials. The permuted accuracy maps from all subjects were combined to create a pool of permuted maps. Then 100,000 group-average accuracy maps were computed by randomly sampling from the pool of permuted maps with replacement to construct a null distribution of accuracy values. These bootstrapped average maps were then thresholded at $p = 0.01$ per voxel and were used for cluster-forming and for constructing the null distribution of cluster sizes for testing the significance of the real group-average map's clusters. Significance of the test was determined at $p < 0.05$ across clusters of size larger than 30 voxels after correcting for multiple comparisons using the FDR [56]. The same set of searchlight analyses were performed where the grand mean of each stimulus condition was removed within each searchlight following the same cluster-based permutation tests and multiple comparison correction methods described above. Results were projected to the cortical surface reconstructed from the Talairach template [55].

#### *Representational similaritiy analysis (RSA)*

To examine the neural representational geometry of the stimulus conditions, we also conducted a representational similarity analysis (RSA) [17]. This was done by calculating the Euclidean distance between patterns of responses for different stimulus conditions. For each ROI, a dissimilarity representational matrix (1-similarity) for the five stimulus conditions was derived. To test whether the similarity between stimulus conditions was mainly driven by physical paths, we conducted a Pearson correlation test between each of these dissimilarity matrices and a hypothesized correlation pattern, which assumed the stimulus conditions with physically different motion paths produced strongest dissimilarity (i.e., 1-similarity = 1) and those with physically same but perceptually different motion paths (i.e., double-drift stimuli) were least dissimilar (i.e., 1-similarity = 0). P values from the correlation analysis were then corrected for multiple tests using FDR [56].

## DATA AND CODE AVAILABILITY

The behavioral dataset and codes generated during this study are available on GitHub at https://github.com/liusr0423/double-drift-fMRI. Neural dataset and codes have not been deposited in a public repository due to file size but are available from the Lead Contact on request.

**Current Biology, Volume *29***

## Supplemental Information

# Neural Correlates of the Conscious Perception of Visual Location Lie Outside Visual Cortex

**Sirui Liu, Qing Yu, Peter U. Tse, and Patrick Cavanagh**

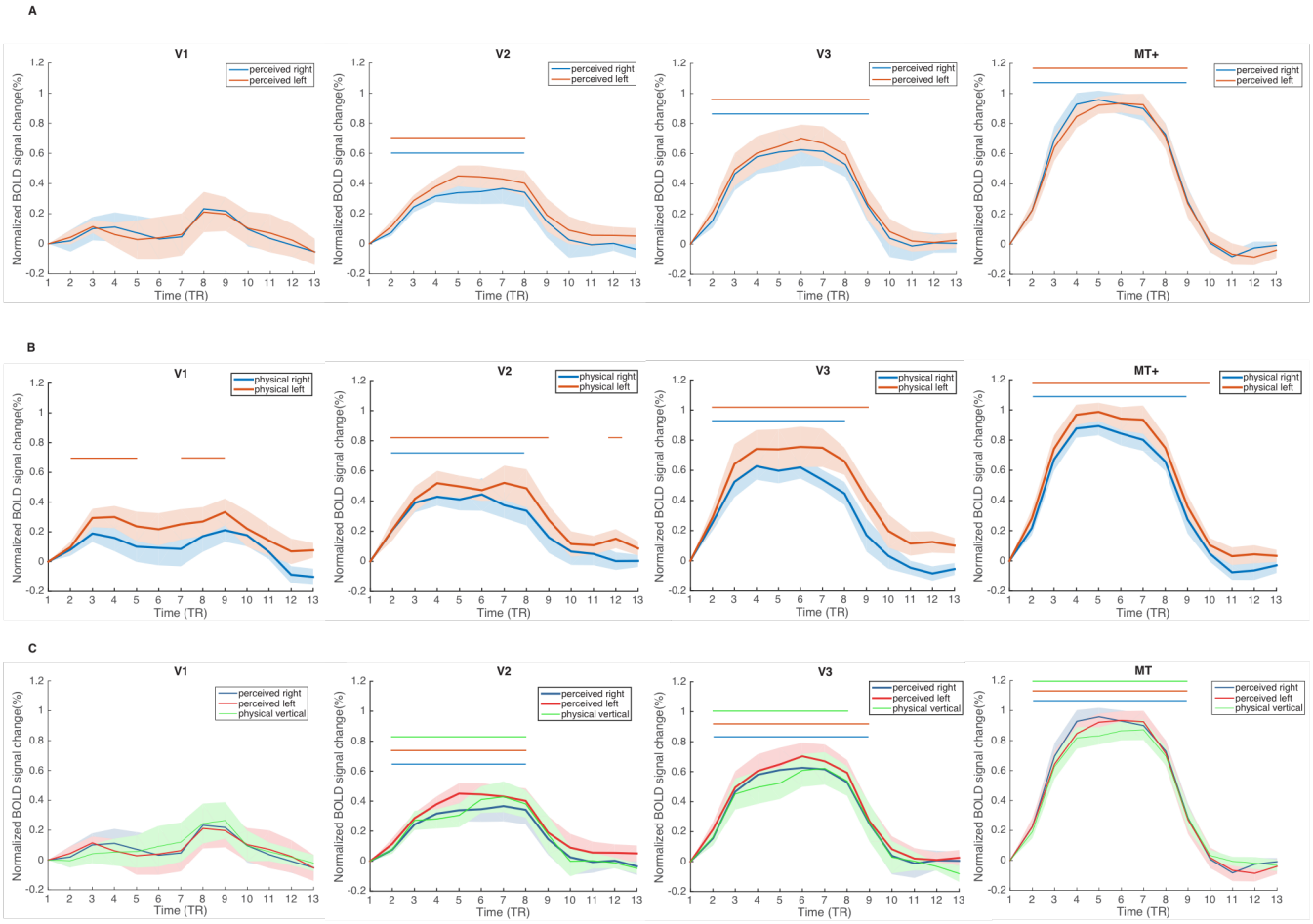


**Figure S1. Experiment 1: mean BOLD time courses of each MPROI (V1, V2, V3 and MT+). Related to Figure 2.**

(A) Mean BOLD time courses for the two double-drift stimuli with perceptually different (i.e. rightward vs. leftward) but physically identical motion directions (i.e. vertical).

(B) Mean BOLD time courses for the two control stimuli with physically rightward vs. leftward motion direction that matched the perceived directions of the double-drift stimuli.

(C) Mean BOLD time courses for the two double-drift stimuli with physically vertical motion path but with internal motions that drove the perceived paths to appear leftward or rightward and for the vertically moving control stimulus with no internal motion.

Error bars represent $\pm$ 1 SEM, horizontal lines at the top of each figure represent time points with significant above-baseline activity for each stimulus condition ($p < 0.05$). Paired samples t-tests showed no significant differences between the two double-drift stimulus conditions, the two control stimulus conditions, or between the double-drift stimuli (with internal motion) and the vertical control stimulus (with no internal motion) that shared the same physical (i.e. vertical) but different perceived motion direction (i.e. illusory left or right direction) ($ps > 0.1$).

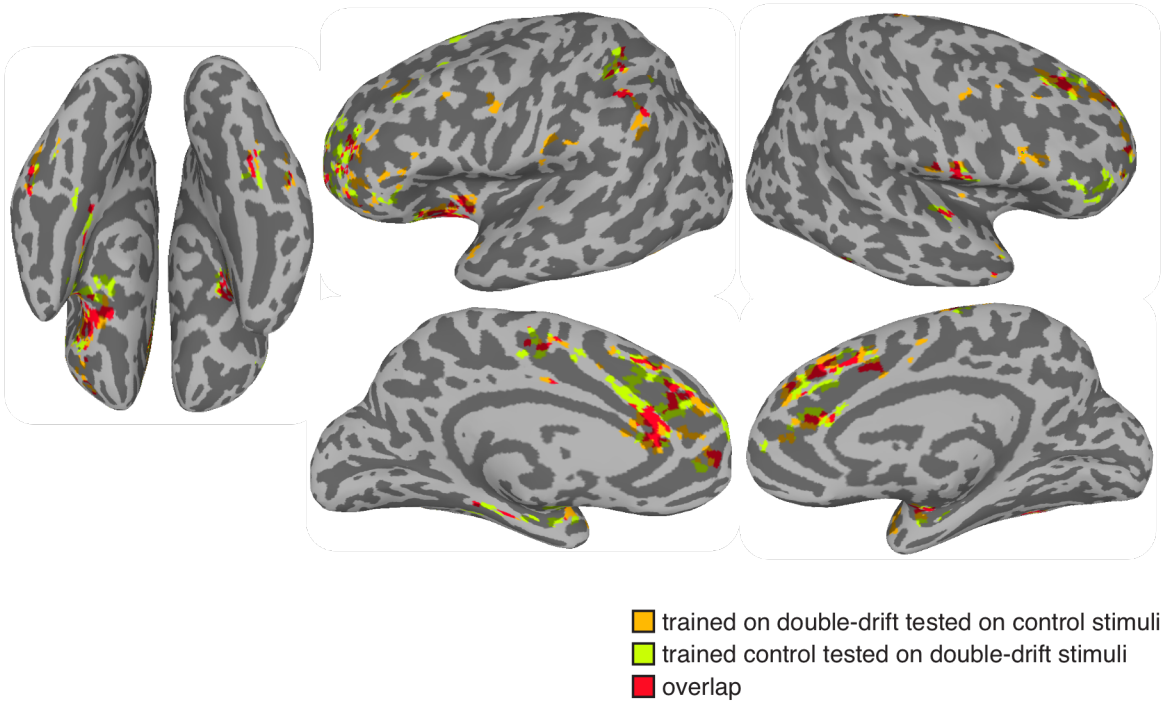


**Figure S2. Experiment 1: cluster-thresholded searchlight map with significant above-chance cross-decoding accuracy when the grand mean of each condition was removed within each searchlight. Related to Figure 4.**

Orange represents significant clusters when training the classifier on double-drift and testing on the control stimuli. Green represents significant clusters when training the classifier on the control and testing on the double-drift stimuli. Results were thresholded at $p = 0.01$ and FDR-corrected across clusters at $p < 0.05$.

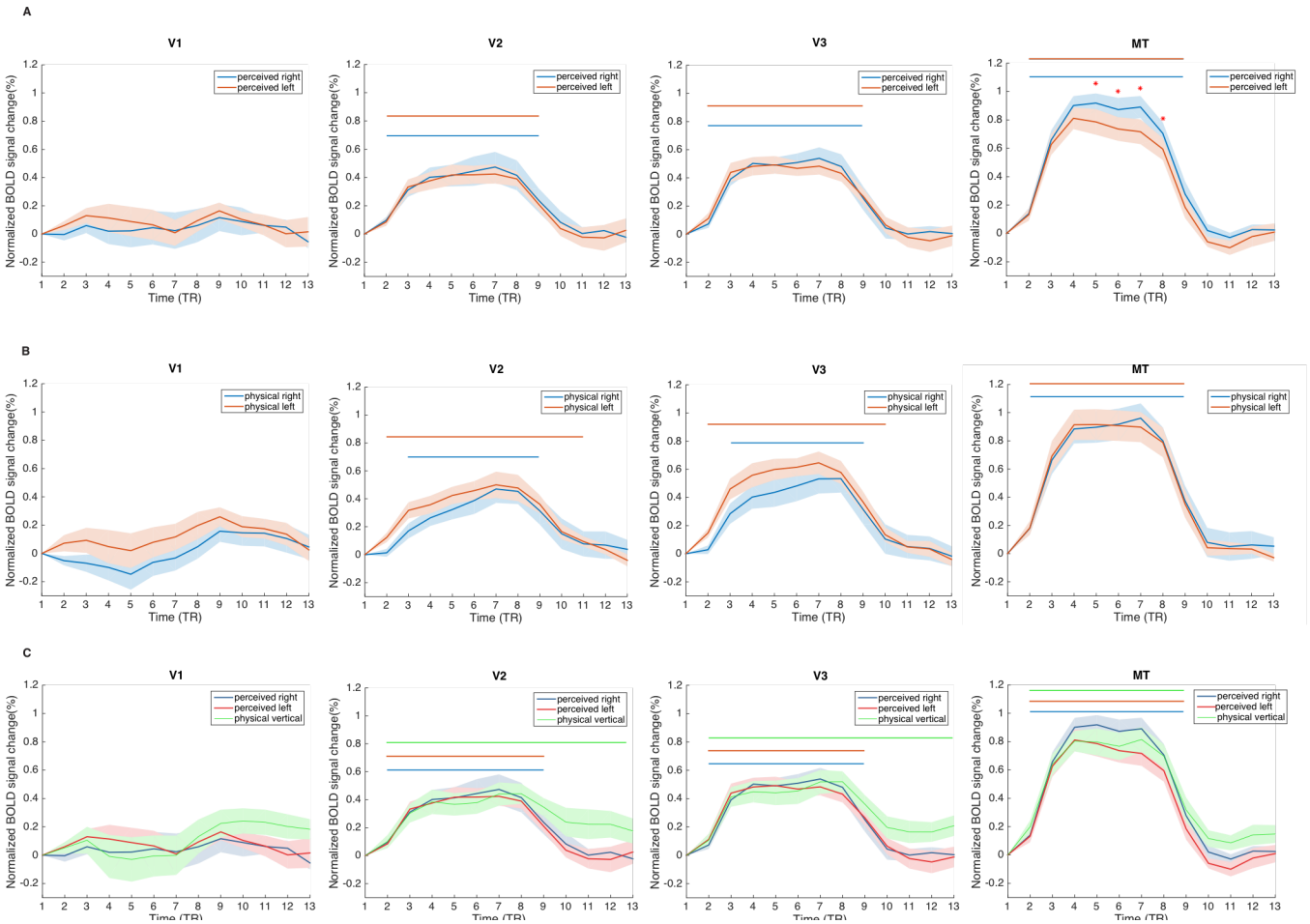


**Figure S3. Experiment 2: mean BOLD time courses of each MPROI (V1, V2, V3 and MT+). Related to Figure 5.**

(A) Mean BOLD time courses for the two double-drift stimuli with perceptually different (i.e. rightward vs. leftward) but physically identical motion directions (i.e. vertical).

(B) Mean BOLD time courses for the two counterphase-flickering control stimuli with physically rightward vs. leftward motion direction that matched the perceived directions of the double-drift stimuli.

(C) Mean BOLD time courses for the two double-drift stimuli with physically vertical motion path but with internal motions that drove the perceived paths to appear leftward or rightward and for the vertically moving counterphase-flickering control stimulus.

Error bars represent $\pm$ 1 SEM, horizontal lines at the top of each figure represent time points with significant above-baseline activity for each stimulus condition ($p < 0.05$). Paired samples t-tests showed no significant differences between the two control stimulus conditions or between the double-drift stimuli (with internal motion) and vertical control stimulus (with counterphase flicker) that shared the same physical (i.e. vertical) but different perceived motion direction (i.e. illusory left or right direction) ($p$s > 0.1). No significant difference between the two double-drift stimulus conditions was observed except in MT+ from TR 5-8 (indicated by red *).

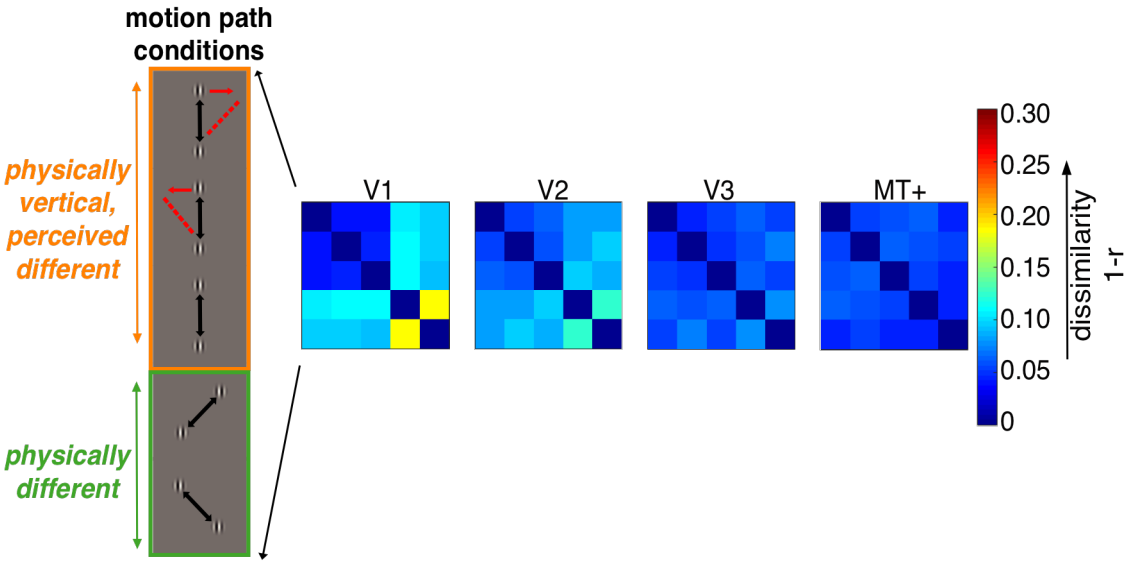


**Figure S4. Experiment 2: representational similarity analysis. Related to Figure 3.**
Representational dissimilarity matrices for the five stimulus conditions in MPROIs (V1, V2, V3 and MT+).

| ROI Name | Size (mean voxel number ± SEM) | Decoding condition | Mean decoding accuracy ± SEM (%) | FDR-adjusted *p* |
|---|---|---|---|---|
| **Experiment 1** | | | | |
| V1 | 72 ± 14 | Illusory paths | 50.6 ± 5.6 | 0.435 |
| | | Physical paths | 90.6 ± 4.4 | 0 |
| | | Illusory to matched physical paths | 45 ± 4.1 | 0.93 |
| | | Matched physical to illusory paths | 47.8 ± 1.9 | 0.73 |
| V2 | 112 ± 22 | Illusory path | 60 ± 2.6 | 0.004 |
| | | Physical paths | 86.7 ± 5.1 | 0 |
| | | Illusory to matched physical paths | 53.3 ± 5.7 | 0.82 |
| | | Matched physical to illusory paths | 50.6 ± 4.9 | 0.73 |
| V3 | 141 ± 70 | Illusory path | 60 ± 3.7 | 0 |
| | | Physical paths | 78.9 ± 4.3 | 0 |
| | | Illusory to matched physical paths | 51.1 ± 7.9 | 0.82 |
| | | Matched physical to illusory paths | 48.9 ± 5.1 | 0.73 |
| MT+ | 194 ± 71 | Illusory path | 53.9 ± 2.3 | 0.12 |
| | | Physical paths | 62.2 ± 2.8 | 0 |
| | | Illusory to matched physical paths | 52.2 ± 3.2 | 0.82 |
| | | Matched physical to illusory paths | 56.7 ± 3.7 | 0.19 |
| **Experiment 2** | | | | |
| V1 | 75 ± 6 | Illusory paths | 52.5 ± 3.2 | 0.157 |
| | | Physical paths | 94.5 ± 2.3 | 0 |
| | | Illusory to matched physical paths | 54 ± 2.2 | 0.420 |
| | | Matched physical to illusory paths | 53 ± 2.5 | 0.668 |
| V2 | 112 ± 6 | Illusory path | 55 ± 3.7 | 0.095 |
| | | Physical paths | 86 ± 3.6 | 0 |
| | | Illusory to matched physical paths | 48 ± 3.8 | 0.729 |
| | | Matched physical to illusory paths | 50.5 ± 3.7 | 0.836 |
| V3 | 106 ± 11 | Illusory path | 61.5 ± 4.1 | 0 |
| | | Physical paths | 72.5 ± 3.7 | 0 |
| | | Illusory to matched physical paths | 49.5 ± 3.1 | 0.704 |
| | | Matched physical to illusory paths | 45.5 ± 2.4 | 0.896 |
| MT+ | 165 ± 20 | Illusory path | 58.5 ± 4.5 | 0.006 |
| | | Physical paths | 58 ± 3.1 | 0.004 |
| | | Illusory to matched physical paths | 51 ± 1.9 | 0.650 |
| | | Matched physical to illusory paths | 47.5 ± 3 | 0.896 |
| **Both experiments** | | | | |
| V1 | | Illusory path | 51.6 ± 3 | 0.22 |
| V2 | | | 57.4 ± 2.3 | 0.001 |
| V3 | | | 60.8 ± 2.7 | 0 |
| MT+ | | | 56.3 ± 2.6 | 0.001 |

**Table S1. Decoding performance and statistical results in each MPROI for Experiment 1, Experiment 2 and two experiments combined. Related to Figure 2 and Figure 5.**

| Decoding condition | Size (voxels) | Peak | | | | areas |
|---|---|---|---|---|---|---|
| | | MNI coordinates | | | Classification accuracy (%) | |
| | | X | Y | Z | | |
| Illusory path | 4720 | 13.5 | -31.5 | 2.5 | 73.3 | Superior frontal gyrus, medial frontal gyrus, anterior cingulate, right parahippocampal gyrus, right superior temporal gyrus |
| | 2993 | -10.5 | 58.5 | -12.5 | 72.2 | Middle occipital gyri, right precuneus, right parahippocampal gyrus, right angular gyrus |
| | 1713 | 22.5 | 10.5 | 29.5 | 72.2 | Left cingulate gyrus, left postcentral gyrus |
| | 294 | -19.5 | 19.5 | 29.5 | 68.9 | Right cingulate gyrus |
| | 220 | 34.5 | 1.5 | -21.5 | 71.1 | Left parahippocampal gyrus, left superior temporal gyrus |
| Matched physical path | 21761 | 13.5 | 85.5 | 8.5 | 84.4 | Left middle occipital gyrus, superior and middle temporal gyri, left inferior parietal lobule, superior and middle frontal gyri, precentral gyrus, postcentral gyrus, paracentral lobule, cingulate gyrus |
| Trained on double-drift tested on control stimuli | 1670 | 7.5 | -34.5 | 26.5 | 67.2 | right superior frontal gyrus, medial frontal gyrus, anterior cingulate |
| | 441 | 4.5 | 22.5 | 2.5 | 65 | Left thalamus |
| | 204 | 31.5 | 46.5 | 26.5 | 65.6 | Left inferior parietal lobule |
| | 168 | -49.5 | -1.5 | 47.5 | 64.4 | Right precentral gyrus, right middle frontal gyrus |
| | 120 | 43.5 | -1.5 | 35.5 | 63.3 | Left precentral gyrus, left middle frontal gyrus |
| | 86 | -40.5 | -28.5 | -3.5 | 65 | Right inferior frontal gyrus |
| | 78 | 37.5 | 46.5 | -24.5 | 64 | Left cerebellum |
| | 48 | 40.5 | 64.5 | 29.5 | 61.7 | Left angular gyrus |
| Trained on control tested on double-drift stimuli | 603 | -7.5 | -22.5 | 26.5 | 65 | medial frontal gyrus, anterior cingulate |
| | 148 | 34.5 | -49.5 | 23.5 | 62.8 | Left superior frontal gyrus |
| | 134 | 25.5 | 37.5 | 47.5 | 63.3 | Left postcentral gyrus |
| | 131 | 7.5 | 28.5 | -3.5 | 63.9 | Left thalamus |
| | 60 | 22.5 | 31.5 | -21.5 | 61.1 | Left cerebellum |
| | 50 | -37.5 | -4.5 | -21.5 | 62.8 | Right superior temporal gyrus, right parahippocampal gyrus |
| | 49 | -61.5 | 10.5 | 20.5 | 61.7 | Right precentral gyrus |
| | 47 | -43.5 | -37.5 | -3.5 | 61.7 | Right inferior frontal gyrus |
| | 36 | 22.5 | -1.5 | -12.5 | 61.1 | Left parahippocampal gyrus |
| | 33 | -13.5 | 19.5 | 59.5 | 61.7 | Right precentral gyrus |
| | 30 | -4.5 | -52.5 | 23.5 | 60 | Right superior frontal gyrus |

**Table S2. Experiment 1: significant clusters found in within- and cross-condition classification searchlight analyses (thresholded at $p = 0.01$ and FDR-corrected across clusters at $p < 0.05$). Related to Figure 4.**

| Decoding condition | Size (voxels) | Peak | | | | areas |
|---|---|---|---|---|---|---|
| | | MNI coordinates | | | Classification accuracy (%) | |
| | | X | Y | Z | | |
| Trained on double-drift tested on control stimuli | 517 | -19.5 | -46.4 | 23.5 | 64.4 | right superior frontal gyrus, right medial frontal gyrus |
| | 402 | 19.5 | -46.5 | -0.5 | 65.6 | Left anterior cingulate, left superior frontal gyrus |
| | 185 | 31.5 | 40.5 | 35.5 | 64.4 | Left inferior parietal lobule |
| | 145 | -25.5 | 4.5 | -15.5 | 61.7 | Right parahippocampal gyrus |
| | 119 | -40.5 | 49.5 | -36.5 | 63.9 | Right cerebellum |
| | 89 | -10.5 | 13.5 | 65.5 | 64.4 | Right superior frontal gyrus, right medial frontal gyrus |
| | 82 | 4.5 | 31.5 | 2.5 | 64.4 | Left thalamus |
| | 75 | 46.5 | 49.5 | -24.5 | 66.1 | Left cerebellum |
| | 75 | -49.5 | -13.5 | 41.5 | 62.2 | Right middle frontal gyrus |
| | 46 | -10.5 | -31.5 | 8.5 | 61.7 | Right anterior cingulate |
| | 45 | -16.5 | -37.5 | 5.5 | 61.7 | Right anterior cingulate, right medial frontal gyrus |
| | 44 | -61.5 | 10.5 | 20.5 | 61.7 | Right postcentral gyrus |
| | 34 | 46.5 | -4.5 | 32.5 | 61.7 | Left inferior frontal gyrus, left precentral gyrus |
| | 31 | 10.5 | 19.5 | -12.5 | 62.2 | Left parahippocample gyrus |
| Trained on control tested on double-drift stimuli | 504 | 34.5 | -52.5 | 11.5 | 66.1 | Left middle frontal gyrus, left superior frontal gyrus |
| | 487 | -13.5 | -28.5 | 38.5 | 65.6 | Right medial frontal gyrus |
| | 161 | 13.5 | 25.5 | -12.5 | 63.9 | Left cerebellum |
| | 121 | 28.5 | 43.5 | 32.5 | 63.9 | Left inferior parietal lobule |
| | 112 | 22.5 | -19.5 | -9.5 | 63.3 | Left inferior frontal gyrus |
| | 108 | -49.5 | -34.5 | 11.5 | 63.3 | Right inferior frontal gyrus |
| | 81 | -25.5 | 1.5 | -15.5 | 64.4 | Right parahippocampal gyrus |
| | 53 | -22.5 | -52.5 | 14.5 | 62.2 | Right superior frontal gyrus |
| | 48 | 22.5 | 13.5 | -15.5 | 61.7 | Left parahippocampal gyrus |
| | 42 | 52.5 | 43.5 | 41.5 | 64.4 | Left inferior parietal lobule |
| | 41 | -10.5 | 16.5 | 59.5 | 65.6 | Right medial frontal gyrus |
| | 38 | -58.5 | 7.5 | 17.5 | 63.3 | Right postcentral gyrus |
| | 33 | 10.5 | -13.5 | 56.5 | 61.7 | Left superior frontal gyrus |
| | 32 | 7.5 | 10.5 | 44.5 | 61.1 | Left paracentral lobule |
| | 30 | -31.5 | 37.5 | -18.5 | 61.1 | Right cerebellum |

**Table S3. Experiment 1: significant clusters found in cross-decoding searchlight analysis when the grand mean of each condition was removed within each searchlight (thresholded at $p = 0.01$ and FDR-corrected across clusters at $p < 0.05$). Related to Figure 4.**

| Decoding condition | Size (voxels) | Peak | | | | areas |
|---|---|---|---|---|---|---|
| | | MNI coordinates | | | Classification accuracy (%) | |
| | | X | Y | Z | | |
| Illusory path | 14543 | -28.5 | -31.5 | 2.5 | 75.5 | Left middle and inferior frontal gyrus, left precentral gyrus, left superior and middle temporal gyrus, left paracentral lobule, left parahippocampal gyrus, right superior and middle frontal gyrus, right superior temporal gyrus, right fusiform gyrus, inferior parietal lobules, postcentral gyrus, anterior cingulate gyrus, supramarginal gyrus |
| Matched physical path | 7533 | 10.5 | 76.5 | -12.5 | 80 | Left superior and middle frontal gyrus, left supramarginal gyrus, left middle temporal gyrus, left precuneus, left calcarine gyrus, left paracentral lobule, right middle frontal gyrus, right middle occipital gyrus, right anterior cingulate gyrus |
| | 222 | -49.5 | 43.5 | 38.5 | 65.5 | Right inferior parietal lobule |
| Trained on double-drift tested on control stimuli | 379 | 34.5 | -31.5 | 35.5 | 64.5 | Left middle frontal gyrus |
| | 166 | -13.5 | 46.5 | 65.5 | 62.5 | Right postcentral gyrus |
| | 79 | -34.5 | -34.5 | 26.5 | 61.5 | Right middle frontal gyrus |
| | 71 | 37.5 | 70.5 | 26.5 | 59.5 | Left angular gyrus |
| | 44 | -49.5 | 4.5 | -24.5 | 61 | Right fusiform gyrus |
| | 43 | -28.5 | -22.5 | 38.5 | 63.5 | Right middle frontal gyrus |
| | 33 | 13.5 | 94.5 | 14.5 | 60.5 | Left middle occipital gyrus |
| Trained on control tested on double-drift stimuli | 395 | -13.5 | 19.5 | 50.5 | 64 | Right medial frontal gyrus |
| | 80 | -40.5 | -34.5 | 35.5 | 62.5 | Right middle frontal gyrus |
| | 77 | -58.5 | -4.5 | -15.5 | 62 | Right middle temporal gyrus |
| | 72 | 1.5 | -1.5 | 41.5 | 63 | Left cingulate gyrus |
| | 59 | 22.5 | -28.5 | 38.5 | 62 | Left middle frontal gyrus |
| | 56 | 46.5 | 34.5 | 14.5 | 60.5 | Left superior temporal gyrus |
| | 36 | 37.5 | -13.5 | 41.5 | 61 | Left middle frontal gyrus |
| | 35 | 58.5 | 7.5 | 20.5 | 60 | Left postcentral gyrus |

**Table S4. Experiment 2: significant clusters found in the within- and cross-condition classification searchlight analyses (thresholded at $p = 0.01$ and FDR-corrected across clusters at $p < 0.05$). Related to Figure 6.**